Gate recurrent units(GRU)

As stated a GRU model is a type of RNN model. GRU's can be compared to a more efficient and simplified way of the LSTM model. To solve the gradient problem of a standard RNN, a GRU uses a reset and update gate. A GRU also has no cell state and uses the hidden state to transfer the 'Opschaler data'.

As mentioned above a GRU only has two gates: a update gate and a reset gate. These vectors decide what information is going to be passed to the output. A GRU is 'special' because it can be trained to keep information from a long time ago without removing relevant information. This makes the GRU model more suitable for predicational methods.

The update gate helps determine how many of the previous information it needs to make a prediction. That is really positive because the GRU can decide to copy all the information from the past steps and eliminate the risk of vanishing gradient problem. The reset gate however is a gate that decides how much past steps it can forget. In short the advantages entail:

- The GRU is faster than an LSTM due to fewer tensor operations.
- The GRU has total control over the flow of information without using a memory unit.
- The GRU model works more efficiently compared to an LSTM model, due to its more efficient use of gates.

Disadvantages±

- An LSTM can use more parameters than a GRU
- Similarly there is issue of increasing gradients at each step called as exploding gradients